GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:49:36 ; Search time 140.11 Seconds
                 (without alignments)
                 80.102 Million cell updates/sec

Title: US-09-331-631A-5
Perfect score: 3326
Sequence: 1 QCMQLETSGQMRRCVSQCDK.........SPRSTKQQQPLVSILDFVGF 625

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
           1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
           2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
           3: /cgn2_6/ptodata/2/iaa/5_COMB.pep:*
           4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
           5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1116	33.6	566	1	US-07-955-905A-2	Sequence 2, Appli
2	1116	33.6	566	1	US-07-955-905A-22	Sequence 22, Appl
3	1060	31.9	587	1	US-07-955-905A-23	Sequence 23, Appl
4	857.5	25.8	605	1	US-07-955-905A-24	Sequence 24, Appl
5	833	25.0	571	1	US-07-955-905A-25	Sequence 25, Appl
6	712.5	21.4	410	1	US-07-955-905A-26	Sequence 26, Appl
7	552	16.6	421	1	US-07-955-905A-27	Sequence 27, Appl
8	216	6.5	1898	1	US-08-056-200-94	Sequence 94, Appl
9	216	6.5	1898	2	US-08-728-323A-2	Sequence 94, Appl
10	216	6.0	1162	2	US-08-800-644-94	Sequence 94, Appl
11	199.5	6.0	683	1	US-07-853-913-2	Sequence 2, Appli
12	143.5	4.3	1805	1	US-08-685-871-2	Sequence 2, Appli
13	139	4.2	1354	3	US-09-234-613-47	Sequence 47, Appl
14	138	4.1	432	3	US-08-918-914-4	Sequence 2, Appli
15	138	4.1	432	2	US-08-285-440-6	Sequence 4, Appli
16	138	4.1	788	2	US-08-630-349-6	Sequence 6, Appli
17	136.5	4.1	558	1	US-08-285-440-6	Sequence 4, Appli
18	136.5	4.1	558	1	PCT-US95-03236-21	Sequence 6, Appli
19	136.5	4.1	614	1	5210183-3	Sequence 21, Appl
20	129	4.0	186	2	US-08-557-309B-43	Patent No. 5210183
21	129	3.9	186	3	US-08-834-306-43	Sequence 43, Appl
22	122.5	3.7	1001	1	US-09-060-410-2	Sequence 2, Appli
23	121	3.6	2101	1	US-08-195-487-4	Sequence 43, Appl
24	121	3.6	2101	4	PCT-US93-06160-4	Sequence 4, Appli
25	120.5	3.6	3248	4	US-08-353-700-1	Sequence 2, Appli
26	120.5	3.6	3248	2	PCT-US95-16216-1	Sequence 1, Appli
27	120	3.6	514	2	US-08-960-022-14	Sequence 4, Appli
28	120	3.6	930	3	US-09-283-763-2	Sequence 2, Appli
29	120	3.6	2101	1	US-08-466-390-4	Sequence 3, Appli
30	120	3.6	2101	1	US-08-470-950-4	Sequence 4, Appli
31	120	3.6	2101	1	US-08-467-781-4	Sequence 4, Appli
32	120	3.6	2101	2	US-08-483-924-4	Sequence 4, Appli
33	115.5	3.5	344	3	Patent No. 5210183-2	Patent No. 521018
34	115	3.5	714	2	US-08-990-114-3	Sequence 5, Appli
35	114.5	3.4	303	1	US-08-109-391A-2	Sequence 2, Appli
36	114.5	3.4	303	1	US-08-630-349-5	Sequence 5, Appli
37	114.5	3.4	303	2	US-08-460-428A-2	Sequence 2, Appli
38	114.5	3.4	1227	2	US-08-459-019A-2	Sequence 2, Appli
39	113.5	3.4	532	1	US-08-760-075A-18	Sequence 18, Appl
40	113.5	3.4	532	1	US-08-285-440-5	Sequence 5, Appli
41	113.5	3.4	532	2	US-08-475-844-9	Sequence 9, Appli
42	112.5	3.4	727	2	US-08-458-860A-2	Sequence 2, Appli
43	112.5	3.4	727	4	PCT-US95-08429-9	Sequence 9, Appli
44	112.5	3.4	754	3	US-09-214-564A-2	Sequence 2, Appli
45	112	3.4	2482	1	US-08-328-254-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-07-955-905A-2

Sequence 2, Application US/07955905A
Patent No. 5770433
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
  TITLE OF INVENTION: PRECURSOR
  NUMBER OF SEQUENCES: 28
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/955,905A
  FILING DATE: 21-JAN-1993
  CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 566 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
US-07-955-905A-2

Query Match   33.6%; Score 1116; DB 1; Length 566;
Best Local Similarity   40.6%; Pred. No. 3.3e-95;
Matches 234; Conservative 107; Mismatches 178; Indels 58; Gaps 13;

```
Qy  68 ICEEEEEYNR---QRDPQQQYEQCRCKRCQQRRETEPRHMQICQQORCERRYEKERRQOKRY 124
          | |:             |::||||:::|||:||: ||| ||:::|| :|||||||
Db  22 LCSGVSAYGRKQYERDPRQQYEQCQCRRCESEATEEREQCEQCERCER----------EY 70

Qy 125 EEQQOREDEERKYEERMKEGDNKRDPQQREYEDCRRHCEQQE--PRLQYQCQRRCQEQOROH 182
         :|||| ||    ||:         ||:||: ::||:|:::|  |:: |:||||| :||:
Db  71 KEQQORQQEEL------------QROYQOCQGRCQEQOOQORReQOOOCCORKCWEQYKEQ 116

Qy 183 GRGGDLMNPQRGGSGRYEEGEEKQSDNPYYFDE-RSLSTRFRTEEGHISVLENFYGRSKL 241
         |:|| |||:||| ||  |::|::   ||:|  ||||||:|||||||: :||| || |
Db 117 ER-GEHENYHNHKKNRSEEEEGQQRNNPYYFPKRRSFQTFRFRDEEGNFKILQRPAENSPP 175

Qy 242 LRALKNYRLVLLEANPNAEVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIR 301
         |:|  | || |||||| |: |:||  | ||::|||| |||| |||||||:|  |:|: |||
Db 176 LKGINDYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTITFVTHENKESYNVQRGTVVS 235

Qy 302 IPAGTTFYLINRDNNERLHIAKFLQTIISTPGQYKEFFAGGQNPEPYLSTFSKEILEAAL 361
         ||||:|| | |:| :  |  |||  ||:|||||||| ||: |||||||| || |::|
Db 236 VPAGSTVVVSQDNQEKLTIAVLALPVNSPGKYELFFPAGNNKPESYGAFSYEVLETVF 295
```

```
RESULT  2
US-07-955-905A-22
; Sequence 22, Application US/0795905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION:  RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION:  PRECURSOR
; NUMBER OF SEQUENCES:  28
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/07/955,905A
; FILING DATE:  21-JAN-1993
; CLASSIFICATION:  435
; INFORMATION FOR SEQ ID NO:  22:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  566 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
; ORIGINAL SOURCE:
; ORGANISM:  Theobroma cacao
; FEATURE:
; NAME/KEY:  Protein
; LOCATION:  1..566
; OTHER INFORMATION:  /note= "67 kd Precursor Protein"
US-07-955-905A-22

Query Match             33.6%;  Score 1116;  DB 1;  Length 566;
Best Local Similarity   40.6%;  Pred. No. 3.3e-95;
Matches  234;  Conservative  107;  Mismatches 178;  Indels  58;  Gaps  13;

Qy  362 NTQTERLRGVLGQQR-----------EGVIIRASQEQIRELTLTRDDSESRRWHIRRGGESSRGP 413
          | :|  |||| |     |||| ||| ::|  | | |||||: :|||||| | |||  |
Db  296 NTQREKLEEILEEQRGQKROQGQGGMFRKAKPEQIRAISQQATSPR----HRGGE--RLA 349

Qy  414 YNLFNKRPLYSNKYGQAYEVKREDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVV 473
          ||| :  |||| :  |:||| |||| ||| | ||| ||||||  |||| |||| :|||||
Db  350 INLLSQSPVYSNQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIFVPHYNSKATFVVFV 409

Qy  474 ASGEADVEMACPHLSGRHGGRGGGK--RHEEEEV------HYEQVRARLSKREAIVVLA 525
          :||||||  ||||||| | |  |||    ||||       ||| :|  | | ||||| |
Db  410 TDGYGYAQMACPHLSRQSGSQSGROJRREQEEESEEETFGEFQQVKAPLSPGDVFVAPA 469

Qy  526 GHPVVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEE 585
          || | |  ||||| ||| | ||||||||||| ||||  |||  :||||| || || ||||
Db  470 GHAVTFFASKDQPLNAVAFGLNAQNNQRIFLAGK-KNLVRQMDSEAKELSFGVPSKLVDN 528

Qy  586 LFNSQDESIFFPGPRQHQQQSPRSTKQQQPLVSILDF 622
          :|| |||| || ||   |:| | |  ||||||||||
Db  529 IFNNPDESYFMSFSQQRQR---RDERRGNPLASILDF 562
```

```
RESULT  3
US-07-955-905A-23
; Sequence 23, Application US/0795905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION:  RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION:  PRECURSOR
; NUMBER OF SEQUENCES:  28
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/07/955,905A
; FILING DATE:  21-JAN-1993
; CLASSIFICATION:  435
; INFORMATION FOR SEQ ID NO:  23:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  587 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  protein
; ORIGINAL SOURCE:
; ORGANISM:  Gossypium hirsutum
; FEATURE:
; NAME/KEY:  Protein
; LOCATION:  1..587
; OTHER INFORMATION:  /note= "Vicilin from G. hirsutum"
US-07-955-905A-23

Query Match             31.9%;  Score 1060;  DB 1;  Length 587;
Best Local Similarity   38.1%;  Pred. No. 5.5e-90;
Matches  228;  Conservative  116;  Mismatches 181;  Indels  74;  Gaps  14;

Qy  302 IPAGTTFYLINRDNNERLHIAKFLQTISTPGQYXEFFPAGGQNPEPYLSTFSKEILEAAL 361
          :||| ||:|:|| | ||| :|| | |:  ||||  :||||  ||  :|| | |:|||| ||
Db  236 VPAGSTVVVVSQDNQEKLTIAVLALPVNSPGKYELFFPAGNNRPESYGAFSYEVLETVF 295

Qy  362 NTQTERLRGVLGQQR-----------EGVIIRASQEQIRELTLTRDDSESRRWHIRRGGESSRGP 413
          | :|  |||| |     |||| ||| ::|  | | |||||: :|||||| | |||  |
Db  296 NTQREKLEEILEEQRGQKROQGQGGMFRKAKPEQIRAISQQATSPR----HRGGE--RLA 349

Qy  414 YNLFNKRPLYSNKYGQAYEVKREDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVV 473
          ||| :  |||| :  |:||| |||| ||| | ||| ||||||  |||| |||| :|||||
Db  350 INLLSQSPVYSNQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIFVPHYNSKATFVVFV 409

Qy  474 ASGEADVEMACPHLSGRHGGRGGGK--RHEEEEV------HYEQVRARLSKREAIVVLA 525
          :||||||  ||||||| | |  |||    ||||       ||| :|  | | ||||| |
Db  410 TDGYGYAQMACPHLSRQSGSQSGROJRREQEEESEEETFGEFQQVKAPLSPGDVFVAPA 469

Qy  526 GHPVVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEE 585
          || | |  ||||| ||| | ||||||||||| ||||  |||  :||||| || || ||||
Db  470 GHAVTFFASKDQPLNAVAFGLNAQNNQRIFLAGK-KNLVRQMDSEAKELSFGVPSKLVDN 528

Qy  586 LFNSQDESIFFPGPRQHQQQSPRSTKQQQPLVSILDF 622
          :|| |||| || ||   |:| | |  ||||||||||
Db  529 IFNNPDESYFMSFSQQRQR---RDERRGNPLASILDF 562
```

```
Qy   68 ICEEEEYNR----QRDPQOQYEQCQCKRCQRRETEPRIHMQICOQRCERRYEKEKRKQQKRY 124
          :|      ||     |||  |||| |  |   |    ||  |  ||  |:|:|  | |  |
Db   22 LCSGVSAYGRKQYERDPROQYEQCORRCESEATEEROQEQCORCER-----EY 70

Qy  125 EEQQREDEEKYEERMKEGDNKRDPQQORYEDCRRHCEQQE--PRLQYQCQRCQEQOROH 182
          |:|| |  |  ::  | | | :|   |||| |||||||||   |||||  | | || |
Db   71 KEQQROQEEL---------QROYQOCQGRCQCEQQOGOREQOCORCWEQYKEQ 116

Qy  183 GRGGDLMNPQRGGSGRYEEGEEKQSDNPYYFDE-RSLSTRFRTEEGHISVLENFYGRSKL 241
          | |||   |  | :||  |  |:  |||| | |  :|| :|||:|:|  ::|||||| |||
Db  117 ER-GEHENYHHIKKNRSEEEGQQQRRNPYYPPKRRSPQTRFRDEEGNFKILQRFAENSPP 175

Qy  242 LRALKNYRLVLLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIR 301
          | | |:||| | ||||||  | || ||| ||  | |||| | | ||| |||| | || ||
Db  176 LKGINDYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTITFVTHENKESTNVQRGTVVS 235

Qy  154 EDCRRHCEQQE---PRLQYQCQRCQEQOROHGRGGDLMNPQR-------GGSGRYEEGEEK 205
          :::|:::       |||||  | | ||  |  |||   |  ||       ||:||||| |:
```

Db 124 OECQOHCHQQEQRPERKQQCVAECRERYQE------NPWRREREEEAEEETEEGEQE 175
Qy 206 QSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTH 265
Db 176 QSHNPFHFHRRSFQSRFEEHGNPRVLQRFASRHPILRGINEFRLSILEANPNTEVLPHH 235
Qy 266 LDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTPYLINRDNNERLHIAKFL 325
Db 236 CDAEKILVTNGRGTLTFLTHENKESYNVVPAGSTVYLANQDNKEKLIIAVLH 295
Qy 326 QTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLG------QQREG 378
Db 296 RPVNNPRQFEEFFPAGSQRPQSYLRAFSREILEPAFNTRSEQLDEFCGRQSHRRQQOG 355
Qy 379 VIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDY 438
Db 356 MFRKASQEQIRALSQEATSPR------EKSGE--RFAFNLLYRTPRYSNQNGREYEACPREF 409
Qy 439 RQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGK 498
Db 410 RQLSDINVTVSALQLNGSIFVPHYNSKATFVVLVNEGNGYVEMVSPHLPRQSSFEEEEE 469
Qy 499 RHEEEEEV------HYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLLFAFG------IN 547
Db 470 QQQEQCEEERRSGQYRKIRSQLSRGDIFVVPANFPVTFVASQNQNLRMTGFGLYNQNIN 529
Qy 548 AQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQDESIFFPGPRQHQQQS 606
Db 530 PDINQRIFVAGKINHV-ROWDSQAKELAFGVSSRLVDEIFNNNPQESYFVS-RQRQRAS 586

RESULT 4
US-07-955-905A-24
Sequence 24, Application US/07955905A
Patent No. 577O433
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
TITLE OF INVENTION: PRECURSOR
NUMBER OF SEQUENCES: 28
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/955,905A
FILING DATE: 21-JAN-1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Glycine max
FEATURE:
NAME/KEY: Protein
LOCATION: 1..605
OTHER INFORMATION: /note= "Vicilin from G. max"
US-07-955-905A-24

Query Match   25.8%;   Score 857.5;   DB 1;   Length 605;
Best Local Similarity 32.1%;   Pred. No. 3.5e-71;
Matches 203;   Conservative 133;   Mismatches 197;   Indels 99;   Gaps 18;

Qy   31 YDNQEDPQTECOQCQRCRQQESDPRQQYCQRRC-----KEICEEEE------EYNR 77
Db   27 YWEKENPKH--NKCLQSC-NSERDSYRNQACHARCNLLKVEKEECEGEIPRPRPRQHP 83
Qy   78 ORDPQQQYEQCOKRCQR------RETEPRHMQICQQRCERRY-EKEKRKQQKRYEEQQRE 130

Db   84 EREPQQPGEKEEDEDEDEQRPIPFPRPQPQPRQEEEHEQRQEQEQWPRKEEKRGGEKGGSEEEDED 143
Qy 131 DEEKYEERMKEGDNKRDPQQOREYEDCRRHCEQQPPRLQYQCQRRCQOCORRCQQHGRGGDLMN 190
Db 144 EDEQQDER--QFPPPRPPHQKE----ERNEEEEED------EEQQRES------ 179
Qy 191 PQRGGSGRYEEGEEKQ-----SDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRAL 245
Db 180 -------EESEDSELRRHKNKPFLFSNRFETLFKNQYGRIRVLQRFNQRSPQLQNL 230
Qy 246 KNYRLVLLFANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAG 305
Db 231 RDYRILEFNSKPNTLLLPNHADADYLIVILNGTAILSLVNNDDRDSYRLQSGDALRVPSG 290
Qy 306 TTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQT 365
Db 291 TTYVVVNPDDNENLRLITLAIPVNKPGREESFFLSSTEAQOSYLQGFSRNILEASYDTKF 350
Qy 366 ERLRGVL------GQOR--EGVIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPY 414
Db 351 EEINKVLFSREEGGQOGEQRLQESVIVEISKEQIRALSKRAKSSSRKTI----SSEDKPF 406
Qy 415 NLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVA 474
Db 407 NLRSRDPIYSNKLGKFFEITPEKNPQLRLDIFILSIVDMNEGALLLPFHFNSKAIVILVIN 466
Qy 475 SGEADVEMACPHLSGRHGGRGGGKRHEEEEEVHYEQ------VRARLSKREAIVVLAGHP 528
Db 467 EGDANIELV-------------GLKEQQOEQQOEQPLEVRKYRAELSEQDIFVIPAGYP 513
Qy 529 VVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFN 588
Db 514 VVV--NATSNLNFFAIGINAENNQRNFLAGSQDNVISQIPSQVQELAFPGSAQAVEKLLK 571
Qy 589 SQDESIFFPGPRQHQQQSPRSTKQQQPLVSIL 620
Db 572 NQRESYFVDAQPKKKEEGNKGRK--GPLSSIL 601

RESULT 5
US-07-955-905A-25
Sequence 25, Application US/07955905A
Patent No. 577O433
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
TITLE OF INVENTION: PRECURSOR
NUMBER OF SEQUENCES: 28
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/955,905A
FILING DATE: 21-JAN-1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 571 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Pisum sativum
FEATURE:
NAME/KEY: Protein
LOCATION: 1..571
OTHER INFORMATION: /note= "Convicilin from P. sativum"
US-07-955-905A-25

RESULT 6
US-07-955-905A-26
Sequence 26, Application US/0795905A
Patent No. 5770433
GENERAL INFORMATION:

```
Query Match    25.0%;    Score 833;    DB 1;    Length 571;
Best Local Similarity 33.2%;    Pred. No. 6.1e-69;
Matches  208;  Conservative  112;  Mismatches  193;  Indels  114;  Gaps  18;

Qy   26  IDWSKYDNQEDPQTECQQCQRRCRQQESDPRQQQYCQRRCKEICEEEEEYNRQRDPCQQY   85
             :  :                : :                           :
Db   25  VTYANYD--EGSETRVPCQQRERGRQ-----------------EGEKEEKRHGEWRPSY   63

Qy   86  EQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQK-----RYEEQQREDEEKYEERMKE  141
            :        : :       :      ::        :                :
Db   64  E-----KEHHE------EEKQKYRYQREKKEQKEVQPGRGRWEREEDEDEQVEEEWRG  109

Qy  142  GDNKRDPQQRREYEDCRRHCEQQEPRLQYQCQRRCQQROHGRGGDLMNPQRGGSGRYEE  201
            ::|:                |||                              ||
Db  110  SQRREDPEER------------ARLRHREER--TKRDRHH------QREGEEERS  145

Qy  202  GEEKQSDNPYYFEDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFV  261
            |  ::|:|| | |  ||||| |||| :: | ||  |||  ||: :|||||: || ||:|
Db  146  SESQEHRNPFLFKSNKFLTLFENENGHIRRLQRFDKRSDLFENLQNYRLVEYRAKPHTIF  205

Qy  262  LPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHI  321
            || ||||: ||| :: | ||||:||||||||| ||||:|:|||||: ||:|:||:||:|
Db  206  LPQHIDADLILVVLNGKAILTVLSPNDRNSVNLERGDTIKIPAGTTISYLVNQDDEEDLRV  265

Qy  322  AKFLQTISTPCQYKEFFPAGGQNPEEYLSTFSKEILEAALNTQTERLRGVL------  372
            ::|: ::: |  |  | |:||| |||||:|:|||||||:|||:|:: :|||
Db  266  VDFVIPVNRPGKFEAF--GLSENKNQYLRGFSKNILEASLNTKYEYIEKVLLEEQEKKPQ  323

Qy  373  ------GQOREGVIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKRPL  422
                     |||:|:||||:|||| |:::| ||| |:| |||||||:|||| |
Db  324  QLRDRKRTQQGEERD-AIIKVSREQIEELRKLAKSSS----KKSLPSEEFEPFNLRSHKPE  378

Qy  423  YSNKYGQAYEVKPE-DYRQLQDMDVSVFIANITQGSMMGPFNTRSTKVVVVASGEADVE  481
            ||||:| |:|  : :|:|||||:|:| ||:| :  ::  |  :|||: : :|:|| ||:
Db  379  YSNKFGKLFEITPEKKYPQLQDLILVSCVEINKGALMLPHYNSRAIVLLVLNEGKGNLE  438

Qy  482  MACPHLSGRHGGRGGGKRHEEEEEVHYE-------QVRARLSKREAIVVLLAGHPVVFVSSG  535
                                : |||::|||:          :||:|||||||||||||| |||
Db  439  LL------------GLKNEQQEREDKRERNNEVQRYEARLSPGDVVIIPACHPVAISASS  486

Qy  536  NENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQDESIF  595
            | |||  ||||||:|| :||| | ||:||| |||| |||||| :|| || : :| |  ||
Db  487  NLNLL--GFGINAKNNQRNFLSGSDDNVISQIENPVKELTFPGSSQEVNRLIKNQKQSHF  544

Qy  596  FPG-PRQHQQQSPRSTKQQQPLVSILD  621
            :   | ||: ::| |||||| |||:||
Db  545  ASAEPEQKEEESQR--KRSPLSSVLD  568
```

ORIGINAL SOURCE:
  ORGANISM: Pisum sativum
FEATURE:
  NAME/KEY: Protein
  LOCATION: 1..410
  OTHER INFORMATION: /note= "vicilin from P. sativum"
US-07-955-905A-26

APPLICANT:
TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
TITLE OF INVENTION: PRECURSOR
NUMBER OF SEQUENCES: 28
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/955,905A
  FILING DATE: 21-JAN-1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 26:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 410 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein

RESULT 7
US-07-955-905A-27
Sequence 27, Application US/0795905A
Patent No. 5770433
GENERAL INFORMATION:

```
Query Match    21.4%;    Score 712.5;    DB 1;    Length 410;
Best Local Similarity 37.9%;    Pred. No. 5.6e-58;
Matches  151;  Conservative  76;  Mismatches  142;  Indels  29;  Gaps  5;

Qy  208  DNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLD  267
            :||  | :|: || | ||||||| | | |||| :|: ||||||| | : | ||| | |
Db   20  ENPFIFKSNRFQTLYENENGHIRLLQKFDKRSKIFENLQNYRLLEYKSKPHTLFLPQYD   79

Qy  268  ADAILLVIGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKFLQT  327
            || ||:|: | | || : :|:|:|:|||| || |:| || : : |||| :|   : || :
Db   80  ADFILVVLSGKATLTVLKSNDRNSFNLERGDAIKLPACSIAYFANRDDNEEPRVLDLAIP  139

Qy  328  ISTPGQYKEFFPAGGQNPEEYLSTFSKEILEAALNTQTERLRGVLGQQR------  376
            : :|||:: | :  | |: ||||||||:|||||:||  |:|::||:| ||
Db  140  VNKPGQLQSFLLSGTQNQKSSLSGFSKNILEAAFNTNYEEIEKVLLEQQEOPQHRRSLK  199

Qy  377  ------EGVIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKRPLYSNKYG  428
                    |:|:|:|:|||:|||| :::||  : : || |||||||:||||  ||||:|
Db  200  DRRQEINEENVIVKVSRDQIEELSKNAKSSS----KKSVSSESGPFNLRSRNPIYSNKFG  255

Qy  429  QAYEVKPEDYRQLQDMDVSVFIANITQGSMMGPFFNTRSTKVVVVVASGEADVEMACPHLS  488
            :|:|  | : :||||:|: | | :| :::   |  ||||: : : | ||| |:  |  |
Db  256  KFEITPEKNQQLQDLDIFPVNSVDIKVGSLLLPNYNSRAIVIVTVTEGKGDFELVGQR--  313

Qy  489  GRIGGRGGGKRHEEEEEV--HYEQVRARLSKREAIVVLLAGHPVVFVSSGNENLLLFAFGI  546
            | :|| ||||||::|||:   :|:||:|||||||||||||||||| ||| | ||| |||:|
Db  314  NENQGKENDKEEQEEETSKQVOLYRAKLSPGDVFVIPAGHPVAINASSDLNLI--GLGI  371

Qy  547  NAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVE  584
            || ||::||||| |:|||| :|| |||||| |  :||
Db  372  NAENNERNFLAGEEDNVISQVERPVKELAFPGSSHEVD  409
```

ORIGINAL SOURCE:
  ORGANISM: Phaseolus vulgaris
FEATURE:
  NAME/KEY: Protein
  LOCATION: 1..421
  OTHER INFORMATION: /note= "vicilin from P. vulgaris"
US-07-955-905A-27

APPLICANT:
TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
TITLE OF INVENTION: PRECURSOR
NUMBER OF SEQUENCES: 28
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/07/955,905A
  FILING DATE: 21-JAN-1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 27:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 421 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein

; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-056-200-94

Query Match   6.5%; Score 216; DB 1; Length 1898;
Best Local Similarity 18.5%; Pred. No. 6.9e-11;
Matches 122; Conservative 107; Mismatches 210; Indels 222; Gaps 21;

```
Qy    6 ETSGQMRCVSQCDKRFEEDIDWSKYDNQEDPQTECQQCQQCRCRQ---------------     50
Db  454 EQKHEQERREQQRLKREQQEERRDWLK----REEETERHEQERRHQQLKRDQEERRWLK    509

Qy   51 -QESDPRQQYCQCRRCKEICEEEEEYNRQR-DPQQQYEQCQKRCQ-----RRETEPRHMQ    103
Db  510 LEEEEERREQQ--ERREQQLRREQEERRQQRLKRDQEEERLQQQRLRSEQQLRREQEERLEQ  567

Qy  104 ICQQRCERRYEKEKRKQQQKRYEEEQQRED------EEKYEERMKEGDNKRDPQQQRYEYDCR 157
Db  568 LLKREEEKRLEQERRQQLRKREQEERRQQQRLRKREQEQERLKREEVE                627

Qy  158 RHCEQQEPRLQYCQCQRRCQCEQQRQHGRGGDLMNPQRGGSGRYEEGEEKQSDNPYYFDERS  217
Db  628 R--LEQEERDERLKREEPEEEERRH----ELLKSEEQEERRHEQLRREQQE--           672

Qy  218 LSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGG    277
Db  673 -----RREQ----RLKREEEEERL                                        687

Qy  278 RGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNNERL--HIAKFLQTISTPGQYK    335
Db  688 EQRLKREHEEEERREQEELAEEE--------QEQARERLKSRIPKWQWQLESEADAR      734

Qy  336 EFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLGQQQREGVIIRASQEQIRELTRDD  395
Db  735 Q------------SKVLLEAPQAGRAEAPQEQEEKRRRESELQWQEEERAHRQQQE        778

Qy  396 SESRR----WHIRGGESSRGPYNLFNKRPLYSNKYGQ-AYEVKPEDYRQLQDMDVSVFI    450
Db  779 EEQRDRDTFWQWQAEEKSERGRQRLSARPLREQQRERQLRAEERQQOREQRFLPE-----   831

Qy  451 ANITQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKRHEEEEVHYEQ    510
Db  832 ----------EEEKE------QRGRQRREREKELQFLE                         853

Qy  511 VRARLSKREAIVVLAGHPVVFVSSGNENLLFAFGINAQNNHENFLAGRERNVLQQIEPQ    570
Db  854 EEEQLQRRERAQQL------QEEEQGLQEDQOERRQEQRRDQ                      889

Qy  571 AMELAFAASRK----EVEELFNSQDESIFFPGPRQHQQQSPRSTKQQ                613
Db  890 KWRWQLEEERKRRRHTLYAKPALQEQLRKEQQLLQEEEEEL----QREEREKRRRQEQ     943

Qy  614 Q                                                               614
Db  944 E                                                               944
```

RESULT 9
US-08-800-644-94
; Sequence 94, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il

Query Match   16.6%; Score 552; DB 1; Length 421;
Best Local Similarity 33.6%; Pred. No. 4.5e-43;
Matches 143; Conservative 71; Mismatches 156; Indels 56; Gaps 12;

```
Qy  203 EEKQSDNPYYFD--ERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFV   261
Db   30 EEESQDNPFYFNSDNSWNTLFKNQYGHIRVLQRFDQQSKRLQNLEDYRLVEFRSKPETLL    89

Qy  262 LPTHLDADAILLVIGGRGALKMIHRDNRESY------NLECGDVIRIPAGTTFYLINRDN    315
Db   90 LPQQADAELLVVRSGSAILLVVKPDDRREYFFLTSDNPIFSDHQKIPAGTIFYLVNPDP-   149

Qy  316 NERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLGQ-   374
Db  150 KEDLRIIQLAMPVNNP-QIHEFFLSSTEAQQSYLQEFSKHILEASNSKFEEINRVLFEE    208

Qy  375 --QREGVIIRASQEQIRELTRDDSESRRWHIRGGESSRGPYNLFNKRPLYSNKYGQAYE    432
Db  209 EGQQEGVIVINIDSEQIKELSKHAKSSSR--------KSLSKQDNIIGNEFGNLTE       255

Qy  433 VKPEDYRQLQDMDVSYFIANI--TQGSMMGPFNTRSTKVVVVASGEADVEMACPHLSGR    490
Db  256 ---RTDNSLNVLISSIEMEEGALFVPHYYSKAIVILVVNEGEAHVELVGPK-----      303

Qy  491 HGGRGGGKRHEEEEVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLFAFGINAQN    550
Db  304 -----GNK-----ETLEYESYRAELSKDDVFVIPAAYPVAIKATSNVNFT--GFGINANN 351

Qy  551 NHENFLAGRERNVLQQI------EPQAMELAFAASRKEVEELFNSQDESIFFPGPRQHQQQ  605
Db  352 NNRNLLAGKTDNVISSIGRALDGKDVLGLTFSGSGSGDEVMKLINKQSGS-YFVDAHHHQQE 410

Qy  606 SPRSTK   611
Db  411 QQKGRK   416
```

RESULT 8
US-08-056-200-94
; Sequence 94, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Mehods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:

; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of, Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; CURRENT APPLICATION DATA:
; SOFTWARE: Patentin Release #1.0, Version #1.25
; OPERATING SYSTEM: PC-DOS/MS-DOS
; COMPUTER: IBM PC compatible
; MEDIUM TYPE: Floppy disk
; COMPUTER READABLE FORM:
; FILING DATE: 14-FEB-1997
; APPLICATION NUMBER: US/08/800,644
; CURRENT APPLICATION DATA:
; CLASSIFICATION: 424
; FILING DATE: 30-APR-1993
; APPLICATION NUMBER: US 08/056,200
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-800-644-94

Query Match 18.5%; Score 216; DB 2; Length 1898;
Best Local Similarity 6.5%; Pred. No. 6.9e-11;
Matches 122; Conservative 107; Mismatches 210; Indels 222; Gaps 21;

```
Qy    6 ETSGQMRRCVSQCDKRFEEDIDWSKYDNQEDPQTECQOCQORRCRQ                  50
          : ||  :    |  :|                                   :  :
Db  454 EQKHEQERRQRLKREQEERRDWLK----REEETERHEQERRKQQLKRDQEEERRERWLK    509

Qy   51 -QESDPRQQQYCQQRCKEICEEEEEYNRQR-DPQQQYEQCQOKRCQ--------RRETEPRHMQ   103
         |::  ||  |  : ||  ||       ||                     |||  |:
Db  510 LEEEERREQQ--ERREQQLRREQEERREQRLKRQEEERLQQRLRSEQOLRREEQEERLEQ   567

Qy  104 ICQQRCERRYEKEKRKQQKRYEEQQRED-----EEKYEERMKEGDNKRDPQQREYEDCR   157
         : || |  :    :|  |||  :|  :        ||  |
Db  568 LLKREEEKRLEQEERREQRLKREQEERRQORLKRDQEEERLEQRLKREEVE   627

Qy  158 RHCEQQEPRLQYQCQQRCEQQQROHQRGGDLMNPQRGGSGRYEEGEGEBKQSDNPYYFDERS   217
         : |  :   :      :|          :||:  |
Db  628 R--LEQEERRDERLKREEPEEERRH----ELLKSEEQEERRHEQLRREQE-------   672

Qy  218 LSTRFRTEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGG   277
          | :   : |  ||
Db  673 -----RREQ--------RLKREEEERL-------------   687

Qy  278 RGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNNERL--HIAKFLQTISTPGQYK   335
          : |  | |   :    |      :         ||
Db  688 EQORLKREHEEERREQLAEEE------QEQAERIKSRIPKWQWQLESEADAR   734

Qy  336 EFFPAGGQNPEPYLSTFSKEILEAALNTQTERLRGVLGQQREGVIIRASQEQIRELTRDD   395
            |    :  :    | ||      |
Db  735 Q--------------SKVLLEAPQAGRAEAPQQEEKRRESELQWQEEERAHRQQQE   778

Qy  396 SESRR---VHIRRGGESSRGPYNLFNKRPLIYSNKYGQ--AYEVKPEDYRQLQDMDVSVFI   450
          | |:    : |  |    |           |  :|
```

RESULT 10
US-08-728-323A-2
Sequence 2, Application US/08728323A
Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,323A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-728-323A-2

```
Db  779 EQORRDFTWQWQAEEKSERGRQRLSARPPLREQRERQLRAEERQQREQRFLPE------   831
           : |                      |  :  :  :     ||
Qy  451 ANITQGSMMGPFENTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKRHEEEEEVHYEQ   510
Db  832 ------------------------EEEKE-------------QRGRQRERKELQFLE   853
                                  |||
Qy  511 VRARLSKREAIVVLLAGHPVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQ   570
Db  854 EEEQLQRRERAQQL------------------QEEDGLQEDQEERRQEQRRDQ   889
                                              |||
Qy  571 AMELAFAASRK------------EVEELFNSQDESIFFPGPRQHQQQSPRSTKQQ   613
Db  890 KWRWQLEEERKRRRHTLYAKPALQEQLRKEQQLLQEEEEEL------QREEEREKRRQEQ   943
                   |  :   |                   |
Qy  614 Q   614
Db  944 E   944
```

Query Match 6.0%; Score 199.5; DB 2; Length 1162;
Best Local Similarity 23.7%; Pred. No. 1.1e-09;
Matches 50; Conservative 64; Mismatches 86; Indels 11; Gaps 3;

```
Qy   17 QCDKRFEEDIDWSKYDNQEDPQTECQQCQORCRQQESDPRQQQYCQORCKEICEEEEEYN   76
         ::  :: |          : ||   ||   |    ||         |    ||
Db  596 QQDEQOCQODEQQODEQQQDEQQODEQOQDEQQODEQQODEQQODEQQODEQQODEQQ   649
```

```
Qy   77  RQRDPQOQYEQCQQKRCQORRE-TEPRHMQICQQRCERRYEKEKRKQQKRYEQQREDEEKY  135
             ::|:|| ||| ||| || |:|  :|: :|||| ||||:||:||||:|:||:||| | |:|
Db  650  QQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQ  709

Qy  136  EERMKREGDNKRDPQQREYEDCRRHCEQQEPRLQYQCQQRCQEQQORQHGRGGDLMNPQRGG  195
         :|: :| | || ||| ||:|| ||| :||| :| | || |||| |||:|  |: | :||
Db  710  DEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQ---EEQEQQEEQ  765

Qy  196  SGRYEEGEEKQSDNPYYFDERSLSTRFTEE  226
         || :||| | ||: :|| |:| |: :||
Db  766  EQELEEQQEELEDQEQELEEQQEQELEEQQE  796
```

RESULT 11
US-07-853-913-2
; Sequence 2, Application US/07853913
; Patent No. 5338839
; GENERAL INFORMATION:
;   APPLICANT: McKay, Ronald D.G.
;   APPLICANT: Lendahl, Urban
;   TITLE OF INVENTION: Nestin Expression As An Indicator of
;   TITLE OF INVENTION: Neuroepithelial Tumors
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;   STREET: Two Militia Drive
;   CITY: Lexington
;   STATE: Massachusetts
;   COUNTRY: U.S.A.
;   ZIP: 02173
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/853,913
;   FILING DATE: 19920319
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/660,412
;   FILING DATE: 22-FEB-1991
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/603,803
;   FILING DATE: 25-OCT-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/201,762
;   FILING DATE: 02-JUN-1988
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/180,548
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Granahan, Patricia
;   REGISTRATION NUMBER: 32,227
;   REFERENCE/DOCKET NUMBER: MIT-4641AAAA
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-861-6240
;   TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1805 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-07-853-913-2

Query Match   4.38;
Best Local Similarity   17.8%;
Matches 141; Conservative 118; Mismatches 283; Indels 249; Gaps   32;

```
Qy   19  DKRFEEDIDWSKYDNQEDPQTECQQCQRCRQQESDPRQQQYCQRRCKEICEEEEYNRQ  78
             |:|:  |    :| |  |     :|       | | | :|    |    |::|
Db  668  DQMLERLVEKEDQSFPRSPEEDQEACRPLQKENQEPLGYEAAEGQLLERLIEKESQESL  727
```

```
Qy   79  RDPQOQYEQCQQKRCQRRETEPRHMQICQ-QRCERRYEKEKRKQQKRYEQQR------  129
         |:::|| ::|  |:  |:::|  |:||   ||||:|  |:|:  ||:||:
Db  728  RSPEEDQEAGKSLQKENQEPLGYEAAEDQMLERLIEKESQESLKSPEENQRIGKPLERE  787

Qy  130  --------EDEEKY----------------  139
                 ::::||
Db  788  NQKSLRYLEENQETFVPLESRNQRPLRSLEVEEEQRIVKPLEKVSQDSLGSLAEENVQP  847

Qy  140  ----------------KEGDNKRDPQQREYEDCRRHCEQQEPRLQYQC  171
                         || |:: ||::  |: |: |  ::|:
Db  848  LRYLEEDDCINKSLLEDKTHKSLGSLEDRNGDSIIIPQESETQVSLRPPEEEDQRIVNHL  907

Qy  172  QRRCQEQQRQHGRGGDLMNPQRGGSSGRYEG-------------EEKQSDNPYYFDER  216
         :|| |:|:|  ::     |:  ||  ||:              ::|:| ::
Db  908  EKESQELFSRSSEEEEQVM--ERSLEGENHESLSSVEKEDQMVESQLEKESQDSKSLEDE  965

Qy  217  SLSTRFRTEEGHISVLENFYGR----------SKLLRALKNYRLVLLEANPNAFVLPTH  265
         |  |    ::| :  ::     ::|||||| |:
Db  966  SQETFGPLEKENAESLRSLAGQDQEEKLEQEETQQTLRAVGNEQMAVSPPEKVDPELPKP  1025

Qy  266  LDADAILLVIGGRGALKMIHRDNRESY---------NLECGDVIRIPAGTTFYLINR  313
         |  :  ::| | :| :     ::|:|          ||: |::|  | | | :|
Db  1026 LGND--------QEIARSLGKENQESLVSLKEKGIETVKSLET-EIIE-PLETAEEDLER  1075

Qy  314  ----DNNERL---HIAKFLQTISTPGQYKEFFPAGG-----QNEPYLSTFSKELLEEA-  360
             |:|||   ||     :| |:  |  ||:|       ||||  ||| |:::|:
Db  1076 RKSIDTQEPLWSTEVAR--ETVEPP---EDEPPGSLGSVDENRET-LPSLEKESQELSS  1128

Qy  361  ----LNTQTERLRGVLGQQREGVIIRASQEQIRELTRD---DSESRRW----------  401
             |  :  |:| |::|:| ||  |::|::|  |:: ||   |:|:|:|
Db  1129 LGKWNVETRVEDSQQCL-QVEEGLQEEQHQESLREVKQELPSSGCNQQRWEDVVEGKAVGQ  1187

Qy  402  ----HIR-RGGESSRGPYNLFNKRPLYSNKYGQAYEV---KPEDYRQ  440
             |:  ||||  |  |||   |: ::  ::|| |||   ||| ||
Db  1188 EAPLATTGVGTEDKAELILRGQGGEEEAAA----EGELLQDIVGEAWSLGSSEPKEQRV  1242

Qy  441  LQDMDVSVFIANITQGSMMGPFFNTRSTKVVVASGEADVEMACPHLSGRHGGRG-GGKR  499
         |:|::  : |||||||||: |::|  :::|:|:| |  |  ||  | |::|| | |||
Db  1243 PAEA-----LDNLEGGAL----------EVPVAQSMPEVTERDEDRAQAGEQ  1279

Qy  500  HEEEEEVHYEQVRARLSRREAIVVL-----AGHPVFVSSGNENL-LLFAFGINAQNNH  552
         ::| :  | ::|:| |   | ::|:     ||  |:  | ||| ||:  | | ||:|
Db  1280 DSIEVTLGLEAARTGLELEQEVVGLEDPRHFAREEAIPPSLGEESVKAKIAQGLEGPGKE  1339

Qy  553  ENFLAGRERNVL-------QQIEPQAMELAFAASRKEVEE--LFNSQDESIFFPGPRQHQ  603
         |:  ||:|:|||       ::|||   ||| ||:||:||  ||:|:|||:  ||| :|
Db  1340 PKEAGALDSGILELPKTSSEALECQGHEESESMEGWEEEEASLETSDHEGSDAPQPR---  1396

Qy  604  QQSPRSTKQQ  614
         :: | ::|:
Db  1397 ---PPETEEDE  1404
```

RESULT 12
US-08-685-871-2
; Sequence 2, Application US/08685871
; Patent No. 6013499
; GENERAL INFORMATION:
;   APPLICANT: NARUMIYA, Shuh
;   APPLICANT: IWAMATSU, Akihiro
;   TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160
;   NUMBER OF SEQUENCES: 68
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 3000 K Street, N.W., Suite 500
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 2007-5109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible

```
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/685,871
;     FILING DATE: 24-JUL-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 435
;     FILING DATE: 25-JUN-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: JP 8-184102
;     FILING DATE: 14-SEP-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 16887/845
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (202)672-5300
;     TELEFAX: (202)672-5399
;     TELEX: 904136
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1354 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-685-871-2

Query Match          4.2%;     Score 139;  DB 3;  Length 1354;
Best Local Similarity  19.4%;  Pred. No. 0.00057;
Matches  126;  Conservative  102;  Mismatches 238;  Indels 184;  Gaps 27;

Qy     13 RCVSQCDKRFEEDIDWSKYDNQEDPQTECQ---QCQRRCRQQESDPRQQQYCQYQCQRRCKEIC 69
          : ::  :  ::  |  | :|   :  | :::|       ||:  :|:::|::  | || :
Db    415 RTSSNADKSLQESLQKTIYKLEEQLHNEMQLKDEMEQKCRTSNI----------KLDKIMK 465

Qy     70 EEEEYNRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQR 129
          :|     |  |      ||:   | :||  | :   | |::|  ::|    ::
Db    466 ELDEGN-----------QRRNLESTVSQI---------EKEKMLLQHRINEYQR 500

Qy    130 EDEEKYEERMKEGDNKRDPQQOREYEDCRRHCEQQEPRLQYQCQRRCQEQQROHGRGDGL- 188
          | |::| |: || :  |  |    | |   | :  |    :  |    |      || |
Db    501 KAEQENEKR-RNVENEVSTLKDQLEDLKKVSQNSQ-----LANEKLSQLQKQLEEANDLL 554

Qy    189 ----------------MNPQRGGSGRYEEGEEKQSDNPYYFDERSLSTRF 222
                          ||| :    : ::  |  |     ::  :  :|
Db    555 RTESDTAVRLRKSHTEMSKSISQLESLNRELQERNRILENSKSQTDKDYQLQAILEAE- 613

Qy    223 RTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALK 282
          |  || :   | :   :|  |:  || :|:  :      :      :  :  : :|
Db    614 RRDRGH-----DSEMIGDLQ ARITSLQEEVK----------HLKIHLEKVEGERKEAQD 657

Qy    283 MIHRDNRESYNLECGDV-------IRIPAGTTEYLINR----DNNERLHIAKFLQTIST 330
          :|   :: :  |   |       :| |  | |: :    | |:| : |:   | :|
Db    658 MLNHSEKKNNLEI-DLNYKLKSLQQRLEQEVNEHKVTKARLTDKHQSIEEAKSVAMCEM 716

Qy    331 PGQYKEFFPAGQQNPEPYLSTFSKEILEAALN--TQTERLRGVLGQQREGVIIRASQEQI 388
          |  :|  |  || : :|:| |  |   |  |   |  ||| | :  || ::|:|:  ||
Db    717 EKKLKE----------ERFAREKAENRVVQIEKQCSML----DVDLKQSQQKL 755

Qy    389 RELT----RDDSESRRWHIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQLQDM 444
             |     |  :|   | :|| ::  ||    |:   :|  | |  | |    |  |
Db    756 EHLTGNKERMEDEVKNLTLQLEQES---------NKRLLQN-------------EL 790

Qy    445 DVSVFIANITQG--SMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGRGGGKRHEE 502
          || :|   | :  :  | ||    : |::    ::   |   | ::   |   |   ||
Db    791 KTQAFEADNLKGLEKQMKQEINT----LLEAKRLLEFELA--QLTKQYRGNEGQMH-EL 842

Qy    503 EEEVHYEQVRARLSKREAIVVLAGHPVFVSSGNENLLLFAFGINAQNNHENFLAGRERN 562
          :|    :   | :      |   | :||   |  |:   ||| :  |: :  | :|:|
Db    843 QDQLEAEQYFSTLYKTQ-----------VKELKEEI-----------EEKNRENLKKIQE-- 880

Qy    563 VLQQIEPQAMELAFAASRKEVEELFNSQDESIFPGPRQHQQQSPRSTKQ 612
           ||| | |  :   |:|||| ::||| |:| :   | |   |:| ::|
Db    881 LQNEKETLATQLDLAETKAESEQLARGLLEEQYFELTQESKKAASRNQE 930
```

```
Qy    563 VLQQIEPQAMELAFAASRKEVEELFNSQDESIFPGPRQHQQQSPRSTKQ 612

Db    881 LQNEKETLATQLDLAETKAESEQLARGLLEEQYFELTQESKKAASRNQE 930

RESULT 13
US-08-933-750C-47
Sequence 47, Application US/08933750C
Patent No. 5932442

;   GENERAL INFORMATION:
;     APPLICANT: Lal, Preeti
;     APPLICANT: Hillman, Jennifer L.
;     APPLICANT: Bandman, Olga
;     APPLICANT: Shah, Purvi
;     APPLICANT: Au-Young, Janice
;     APPLICANT: Yue, Henry
;     APPLICANT: Guegler, Karl J.
;     APPLICANT: Corley, Neil C.
;   TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
;   NUMBER OF SEQUENCES: 98
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Incyte Pharmaceuticals, Inc.
;     STREET: 3174 Porter Drive
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94304
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/933,750C
;     FILING DATE: September 23, 1997
;     CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Billings, Lucy J.
;     REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0356 US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-855-0555
;     TELEFAX: 415-845-4166
;     TELEX:
;   INFORMATION FOR SEQ ID NO: 47:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 432 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: TLYMNOT04
;       CLONE: 2926777
US-08-933-750C-47

Query Match          4.1%;     Score 138;  DB 2;  Length 432;
Best Local Similarity  22.5%;  Pred. No. 0.00012;
Matches  47;  Conservative  43;  Mismatches 87;  Indels 32;  Gaps 6;

Qy     20 KRFEEDIDWSKYDNQEDPQTECQQCQRRCRQQESDPRQQQYCQYQCQRRCKEICEEEEYNRQR 79
          |: |: : | ||  |||||         :| |   | |   | |     |    :   |  |
Db    236 KRTEEP-----DRDERLKKEKQEEREEREKEREREREREKRR---EEEEREKERAR 285

Qy     80 DPQQQYEQ----------CQKRCQR-----RETEPRHMQICQQRCERRYEKRKQQK 122
          ||  |  |          | |: |      :    ||      |   |  :|  ||
Db    286 DRERRKRSRSRSHSSRPTSDRRCSRSRDIKRSRSRERRTRSRDRRRSRSHDRSERKHRS 345

Qy    123 RYEEQQR------EDEEKYEERMKEGDNKRDPQQOREYEDCRRHCEQQEPRLQYQCQRRCQEQQ 179
          | |:|||      :|  | : :      |  |      | |   |    |    :  |    |
```

Left column (RESULT 14)

```
Db  346 RSRDRRRSKSRDRKSYKHRSKSRDREQDRKSKEKE---KRGSDDKKSSVKSGSREKQSEDT 403
Qy  180 RQHGRGGDLMNPQRGGSGRYEEGGEEKQSD 208
Db  404 NTESKESDTKNEVNGTSEDIKSEGDTQSN 432
```

RESULT 14
US-09-234-613-47
Sequence 47, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TLYMNOT04
CLONE: 2926777
US-09-234-613-47

```
Query Match   4.1%;   Score 138;   DB 3;   Length 432;
Best Local Similarity   22.5%;   Pred. No. 0.00012;
Matches   47;   Conservative   43;   Mismatches   87;   Indels   32;   Gaps   6;

Qy   20 KRFEEDIDWSKYDNQEDPQTECQQCORRCRQOESDPRQOQYCQRRCKEICEEEEEYNRQR 79
Db  236 KRTEEP------DRDERLKKEKQEREREKEREREERKRRR---EEEEREKERAR 285

Qy   80 DPQQQYEQ--------COKRCQR------RETEPRHMQICQQRCCERRYEKERKQOK 122
Db  286 DRERRKRSRSRSRHSSFTSDRRCSRSRDHKRSRSRSRERRTRSRDRRRSRSRHDRSERKHRS 345
```

Right column (RESULT 15)

```
Qy  123 RYEEQOR---EDEERKYEERMKEGDNKRDPQOREYEDCRRHCEQQEPRLQYQCQRRCQOEQQ 179
Db  346 RSRDRRRSKSRDRKSYKHRSKSRDREQDRKSKEKE---KRGSDDKKSSVKSGSREKQSEDT 403
Qy  180 RQHGRGGDLMNPQRGGSGRYEEGGEEKQSD 208
Db  404 NTESKESDTKNEVNGTSEDIKSEGDTQSN 432
```

RESULT 15
US-08-918-914-4
Sequence 4, Application US/08918914
Patent No. 5876963
GENERAL INFORMATION:
APPLICANT: Mitchell, Peter
APPLICANT: Hutchinson, Nancy
APPLICANT: Lawton, Michael
APPLICANT: Magna, Holly
APPLICANT: Yocum, Sue
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,914
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0369
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1070094
US-08-918-914-4

```
Query Match   4.1%;   Score 138;   DB 2;   Length 788;
Best Local Similarity   25.4%;   Pred. No. 0.00031;
Matches   51;   Conservative   32;   Mismatches   68;   Indels   50;   Gaps   7;

Qy   12 RRCVSQCDKRFEEDIDWSKYD-NQEDPQTECQQCQ----------RRCRQOESDPRQ 57
Db  148 RKKRQQDEAVEGVGDWTDWSHCSSNGHEVRSQACEYGRKIQRRGCPARSAPQRVPAPA 207

Qy   58 QQYCQRRCKEICEEEEINRQRDPQOQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEK 117
Db  208 QQYAPR------APEYPSAQQOQQREQQREQOHREHQARLQOHQOQQQQOQQQ 259

Qy  118 RKQOKRYEEQOREDEEKYEERMKEGDNKRDPQQRE-----YEDCRRHCEQQEPRLQYQCQ 172
```

```
Db  260 RPPQPQPQPQPQPPQ---------------RPPQQPQSFSGTHE---LHLQRQREQQQQQQQ 303
            | |:|:|:|                | |||:          | ::| ::|||
Qy  173 RRCQEQQRQHGRGGDLMNPQR 193
        :: |:|||  |||:
Db  304 QQQQQQQRQ--------QNPQQ 317
```

Search completed: March 1, 2001, 15:49:44
Job time: 389 sec